PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The acceptability of and symptom findings from an online symptom check-in tool for COVID-19 outpatient follow-up among a predominantly healthcare worker population.
AUTHORS	Kerr, Colm; O' Regan, Simon; Creagh, Donnacha; Hughes, Gerry; Geary, Una; Colgan, Mary Paul; Canning, Caitriona; Martin, Zenia; Merry, Concepta; Noonan, Noirin; Bergin, Colm

VERSION 1 – REVIEW

REVIEWER	Souroujon, Daniel Tel Aviv University, Public Health
REVIEW RETURNED	25-Apr-2021

GENERAL COMMENTS	BMJopen - Review of "The acceptability of and symptom findings from an online symptom check-in tool for acute COVID-19 outpatient follow-up." 1) Aim (& abstract): Creating the tool is probably not the aim, but the means to arrive at your aim. what was the aim? Improving covid-19 patients' medical outcome? reducing the number of in person triage or readmission? creating higher engagement with outpatients? Focus on more measurable outcomes. Your results and conclusion should correspond with your aims and outcomes. 2) Methods: a. Was any information shown users answering a specific answer during the questionnaire? b. Was there a time frame for the clinician to review the incoming data? 24h? 48h? 3) Results: a. You mention this tool is a triage tool – what was the distribution of the triage? This is one of the most important graphs I would present. b. I would add a more classic Table1 which shows all cohort characteristics like demographics (e.g. age, gender), lifestyle (e.g. smoking habits) & the comorbidities you've described. c. Online check-in tool i. "Fatigue was the most common symptom in the first 4 days of illness, followed by headache and myalgia" - add percentages ii. "The case criteria symptoms of cough, taste & smell disturbance, pyrexia and cough were the 4th, 5th, 6th, 16th and 19th most frequently reported symptoms respectively during this period." - cough appears twice, also, not sure if taste and smell disturbance was counted as one or two symptoms - please rephrase to be clearer. iii. "Most symptoms remained static or decreased over time with the exception of taste, smell disturbance and anorexia which peaked during days 5-8 of symptoms" - again, please add

	quantitative data, from what percentage to what percentage did it increase? d. Feedback Survey i. "16% (23) of respondents to this feedback survey were admitted to hospital during their illness." - during the followup time window or after initial ER visit? do you know the reason for admission (covid related or not?) ii. "Reasons given by those who didn't use the online symptom check-in (n=32) included feeling too unwell to use it, forgetting about it, feeling sufficiently informed by healthcare staff already and feeling better and therefore not having a need to use it." - add distribution numbers and percentages 4) Discussion: a. I would at least discuss the possibility of improving patients' outcome using the tool. But more data would have to be presented for this to be discussed. For example – how many patients were asked to come back to the hospital for readmission because of their symptoms? How many were hospitalized? And was the reason covid related? (if none, was this tool really helpful with dealing with the biphasic nature of the disease…?) b. This was written in the results: "After 6 months only 39% of participants reported being symptom free." - this is very interesting as it differs from the reported percentages - why not discuss this in the discussion? 5) Figures: a. figure 2: this might be portrayed better as a table and not a figure - for your consideration.

REVIEWER	Mills, Brennen Edith Cowan University
REVIEW RETURNED	30-Apr-2021

GENERAL COMMENTS	This paper evaluates an online symptom check-in tool for COVID-19 +ve outpatients from a single hospital. This is a very innovative concept which has a great deal of value. This kind of work is important to get out in the world so others can see and understand the value of such tools. I do have some concerns pertaining to the way data has been collected and interpreted, but I am very much hoping the authors are able to adequately address these comments, as I strongly believe/hope this paper should be approved for publication (pending the addressing of the below comments). Unfortunately, there are some areas of discomfort with the data and its generalisability, particularly with such a large proportion of the sample having a healthcare background. The limitations of such have not been adequately addressed in the current version. Nonetheless, I encourage the authors to address these concerns (which may require a re-examination of some data). ABSTRACT Lines 9-11 – The second sentence is not crystal clear here. Should the word 'that' be removed? Line 44 – Would suggest re-writing or removing the final sentence as it seemingly comes 'out of the blue' and none of the results or the above information from the conclusion discusses or supports this contention. STRENGTHS AND LIMITATIONS

	Line 10-11; Grammatical error ‘among of’ in first dot point. Also, what is wave 1? At this stage of reading the manuscript we don’t have any context other than the Title and the abstract, so authors need to be conscious of this. Line 12; 6-month follow-up from when? The first time they used the app? Following their initial diagnosis? INTRODUCTION Very well written. I would however suggest a stronger link needs to be made between the introduction and aims section to further justify the study. The link is made somewhat, but I think more could be done to demonstrate the need/value of investigating whether an app such as this is worthwhile. Make crystal clear what problem/s specifically are being addressed through the creation and implementation of this app? While it may be obvious to some, it may not be to others, so this needs to be explicitly stated. METHODS Lines 10-14; were there any cases of participants not presenting until after day 3 of their symptoms? If so, did they get a text message at day 4 instead of day 3, or did they just wait until day 6 of their illness? Was the online tool utilised amongst every person diagnosed with COVId-19 from tests being undertaken at St James Hospital during the study period? I understand not everyone may have responded to the text by going through the online portal, but was a text sent to every diagnosed individual, or was there some kind of selection and the sample is not exhaustive? RESULTS Pg 8 Lines 10-28; I note 83% of text recipients were healthcare workers. Is this because the majority of COVID-19 diagnoses from the hospital were healthcare workers? This seems disproportionate, unless the hospital was receiving a far greater proportion of COVID-19 diagnoses from people working in healthcare than people working in other sectors? I also note that 90% of the respondents were healthcare workers. I’m concerned that really you might only be able to generalise the results of this data to healthcare workers who are likely far more invested in providing this kind of information given their background? Overall, it seems from your sample that healthcare workers were far more likely to utilise the tool than non-healthcare workers? Pg 8 Lines 29-31; so the feedback survey was sent to all 413 original recipients, regardless of whether they actually used the online tool or not? How well will those that did not use the tool be able to comment on its efficacy and usability etc. if they never actually used the tool? How many of these140 that responded to the 6-months survey actually utilised the tool at least once? Pg 9 Lines 3-7; How are you getting data from patients 48 days after the onset of their illness if texts were only sent at 3, 6, 9 and 12 days after onset as outlined in your methods? Pg 9 Lines 18-23; What did you classify as case defining criteria and what was this based on? While fatigue may not be case defining, it is certainly inferred as one of the more common symptoms. Figure 3 – would it not be a more worthy comparison to remove the data of participants that you do not have data from 6 months later? Would this not be a truer reflection of illness progression/change over the period if you only included data for those that you have both baseline and follow-up data? Presently, if you include data from participants at baseline, we have no idea

	how their illnesses progressed over time so have no kind of comparison so the data may not be reflective of what actually happened. By removing them, and only focussing on those with both pre and post data, this limitation no longer applies. Pg 9; Lines 38-39; please define what you mean by 'third-level educated?' Pg 9 Lines 42; 21% of participants suggested they were inconvenienced. Did you get any data as to why the felt this way? Do you have data as to how many of these who suggested they were inconvenienced still puckered up and used the tool, or did all of these not use the tool? Pg 9 Line 55-56, so only 108 of the 140 who completed the feedback survey actually utilised the online check-in? Think this needs to be made clear upfront. Pg 9/10 Lines 59-5; The 69% who received a call from the COVID-19 team, was this data collected via self-report from the online questionnaire or was it collected objectively through cross-reference with call logs? Pg 10 Lines 17-18; those that suggested the quality of the service was 8.5 (n=108), wouldn't they others who did not use the tool also be able to comment on the quality of the service as they were still receiving text messages and had the capability to make/receive phone calls? DISCUSSION Pg 10 Lines 54-55; the statistic suggesting that the online check in tool had good uptake/response rate is a little misleading, as yes 72% used the tool at least once, but they had an opportunity to use it every three days for (unclear how long they received texts for). I think you need to be a little more clear with respect to response rates. The statistic you report is valid, but so is the OVERALL response rate from every text that was sent vs. how many of them prompted a login to the online tool to report on symptoms. Limitations – while the authors do a sound job discussing the potential limitations of the online tool, they need to do more to address the limitations associated with the research and evaluation component of the paper. There was a skewed sample consisting primarily of healthcare workers (for example). While we want to know about the tool and if it worked well/not well, you have presented a lot of data you suggest speaks to the efficacy of the tool. The readers need a thorough understanding of the limitations associated with these data so as to inform their own decisions about efficacy.

REVIEWER	Hodgkinson , Ian R. Loughborough University, School of Business and Economics
REVIEW RETURNED	30-Apr-2021

GENERAL COMMENTS	Thank you for the opportunity to review this article titled 'The acceptability of and symptom findings from an online symptom check-in tool for acute COVID-19 outpatient follow-up'. The study presents a novel intervention and offers several important insights. In my comments below I explain how the authors could further strengthen the article and develop more impactful implications and insights from the data: Existing literature: There needs to be some acknowledgement of the growing body of research on telehealth given the centrality of telecommunications technology in the study, for example, see the

	following article: Leite, H., Gruber, T. and Hodgkinson, I.R., 2020. Flattening the infection curve–understanding the role of telehealth in managing COVID-19. Leadership in Health Services, 33(2), 221-226. Data audit: How were the contact details for each patient accessed? There needs to be a statement on the protocol followed at a minimum. Since 90% of respondents were healthcare workers it would be useful to reflect on what this bias in the sample may mean for any implications drawn. For instance, to generalise the implications across the general public would be misleading as the respondents will have a much greater level of knowledge and understanding concerning COVID-19. Follow-up feedback survey: The authors state a feedback survey was used to evaluate acceptability of the online check-in, but this was not sent until 6 months after the SMS 'intervention'. There needs to be a justification for the time lapse as it might be expected that time passed had influenced the perceptions of the participating patients. Findings: the current presentation of findings is highly descriptive and, therefore, offers limited insights into the role of the online tool. A suggestion would be to segment the patients into groups, possibly based on age bands, healthcare role, etc. and statistically examine whether significant differences exist between groups based on their emotional responses to the tool (i.e. reassuring, helpful, worrying, etc.) and whether significant differences exit between groups based on the perceived usefulness of the tool, satisfaction with the tool, quality of the service, and so forth. Currently, the authors only appear to do this for the healthcare worker group and non-healthcare worker group. Without these more fine-grained insights, the study is less impactful and insightful. In the discussion of the findings the authors state "consideration should be given to including symptoms such as fatigue, headache, myalgia, nasal congestion and pharyngitis to the testing criteria." Again, it would be useful to see how these symptoms varied between patient groups and whether certain groups reported greater relative levels of these symptoms. In its current form, these insights are somewhat lost within the text and are not clearly presented in the results.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Comments to the Author:
See attached file.

1) Aim (& abstract): Creating the tool is probably not the aim, but the means to arrive at your aim. what was the aim? Improving covid-19 patients' medical outcome? reducing the number of in person triage or readmission? creating higher engagement with outpatients? Focus on more measurable

outcomes. Your results and conclusion should correspond with your aims and outcomes. Updated, thank you

2) Methods:

a. Was any information shown users answering a specific answer during the

questionnaire? I'm afraid I don't understand this question

b. Was there a time frame for the clinician to review the incoming data? 24h? 48h?

Yes, updated manuscript to reflect. Thank you

3) Results:

a. You mention this tool is a triage tool – what was the distribution of the triage? This

is one of the most important graphs I would present. updated manuscript to reflect. Thank you

b. I would add a more classic Table1 which shows all cohort characteristics like

demographics (e.g. age, gender), lifestyle (e.g. smoking habits) & the comorbidities

you've described. updated manuscript to reflect. Thank you

c. Online check-in tool

i. "Fatigue was the most common symptom in the first 4 days of illness,

followed by headache and myalgia" - add percentages updated manuscript to reflect. Thank you

ii. "The case criteria symptoms of cough, taste & smell disturbance, pyrexia

and cough were the 4th, 5th, 6th, 16th and 19th most frequently reported

symptoms respectively during this period." - cough appears twice, also, not

sure if taste and smell disturbance was counted as one or two symptoms -

please rephrase to be clearer. updated manuscript to reflect. Thank you

iii. "Most symptoms remained static or decreased over time with the exception

of taste, smell disturbance and anorexia which peaked during days 5-8 of

symptoms" - again, please add quantitative data, from what percentage to

what percentage did it increase? updated manuscript to reflect. Thank you

d. Feedback Survey

i. "16% (23) of respondents to this feedback survey were admitted to hospital

during their illness." - during the followup time window or after initial ER

visit? do you know the reason for admission (covid related or not?)

ii. "Reasons given by those who didn't use the online symptom check-in

(n=32) included feeling too unwell to use it, forgetting about it, feeling sufficiently informed by healthcare staff already and feeling better and therefore not having a need to use it." - add distribution numbers and percentages updated manuscript to reflect. Thank you

4) Discussion:

a. I would at least discuss the possibility of improving patients' outcome using the tool. But more data would have to be presented for this to be discussed. For example – how many patients were asked to come back to the hospital for readmission because of their symptoms? How many were hospitalized? And was the reason covid related? (if none, was this tool really helpful with dealing with the biphasic nature of the disease…?) updated manuscript to reflect. Thank you

b. This was written in the results: "After 6 months only 39% of participants reported being symptom free." - this is very interesting as it differs from the reported percentages - why not discuss this in the discussion?

5) Figures:

a. figure 2: this might be portrayed better as a table and not a figure - for your consideration updated manuscript to reflect. Thank you

Reviewer: 2

Comments to the Author:
This paper evaluates an online symptom check-in tool for COVID-19 +ve outpatients from a single hospital. This is a very innovative concept which has a great deal of value. This kind of work is important to get out in the world so others can see and understand the value of such tools. I do have some concerns pertaining to the way data has been collected and interpreted, but I am very much hoping the authors are able to adequately address these comments, as I strongly believe/hope this paper should be approved for publication (pending the addressing of the below comments). Unfortunately, there are some areas of discomfort with the data and its generalisability, particularly with such a large proportion of the sample having a healthcare background. The limitations of such have not been adequately addressed in the current version. Nonetheless, I encourage the authors to address these concerns (which may require a re-examination of some data).

ABSTRACT
Lines 9-11 – The second sentence is not crystal clear here. Should the word 'that' be removed? Corrected, thank you
Line 44 – Would suggest re-writing or removing the final sentence as it seemingly comes 'out of the blue' and none of the results or the above information from the conclusion discusses or supports this contention. Updated, thank you

STRENGTHS AND LIMITATIONS
Line 10-11; Grammatical error ‘among of’ in first dot point. Also, what is wave 1? At this stage of reading the manuscript we don’t have any context other than the Title and the abstract, so authors need to be conscious of this. Corrected, thank you
Line 12; 6-month follow-up from when? The first time they used the app? Following their initial diagnosis? Updated, thank you

INTRODUCTION
Very well written. I would however suggest a stronger link needs to be made between the introduction and aims section to further justify the study. The link is made somewhat, but I think more could be done to demonstrate the need/value of investigating whether an app such as this is worthwhile. Make crystal clear what problem/s specifically are being addressed through the creation and implementation of this app? While it may be obvious to some, it may not be to others, so this needs to be explicitly stated. Updated, thank you

METHODS
Lines 10-14; were there any cases of participants not presenting until after day 3 of their symptoms? If so, did they get a text message at day 4 instead of day 3, or did they just wait until day 6 of their illness? Updated, thank you
Was the online tool utilised amongst every person diagnosed with COVId-19 from tests being undertaken at St James Hospital during the study period? I understand not everyone may have responded to the text by going through the online portal, but was a text sent to every diagnosed individual, or was there some kind of selection and the sample is not exhaustive? Updated, thank you
RESULTS
Pg 8 Lines 10-28; I note 83% of text recipients were healthcare workers. Is this because the majority of COVID-19 diagnoses from the hospital were healthcare workers? This seems disproportionate, unless the hospital was receiving a far greater proportion of COVID-19 diagnoses from people working in healthcare than people working in other sectors? I also note that 90% of the respondents were healthcare workers. I’m concerned that really you might only be able to generalise the results of this data to healthcare workers who are likely far more invested in providing this kind of information given their background? Overall, it seems from your sample that healthcare workers were far more likely to utilise the tool than non-healthcare workers? Majority of those diagnosed were healthcare workers, updated to reflect
Pg 8 Lines 29-31; so the feedback survey was sent to all 413 original recipients, regardless of whether they actually used the online tool or not? How well will those that did not use the tool be able to comment on its efficacy and usability etc. if they never actually used the tool? How many of these140 that responded to the 6-months survey actually utilised the tool at least once? Updated to clarify that only those who used the tool (n=108) gave feedback to this question
Pg 9 Lines 3-7; How are you getting data from patients 48 days after the onset of their illness if texts were only sent at 3, 6, 9 and 12 days after onset as outlined in your methods? Patients could keep accessing and filling in the tool as often as they liked, one patient filled it in 48 days post symptom onset, the vast majority stopped however after receiving their last text message invite (at day 12)

Pg 9 Lines 18-23; What did you classify as case defining criteria and what was this based on? While fatigue may not be case defining, it is certainly inferred as one of the more common symptoms.

Have updated manuscript to emphasise the ECDC case defining clinical criteria of cough, dyspnoea, tase/smell disturbance, shortness of breath

Figure 3 – would it not be a more worthy comparison to remove the data of participants that you do not have data from 6 months later? Would this not be a truer reflection of illness progression/change over the period if you only included data for those that you have both baseline and follow-up data?

Presently, if you include data from participants at baseline, we have no idea how their illnesses progressed over time so have no kind of comparison so the data may not be reflective of what actually happened. By removing them, and only focussing on those with both pre and post data, this limitation no longer applies. Unfortunately the 6 month feedback survey was entirely anonymous so it was not possible to link symptom feedback to individual responses to the check in tool
Pg 9; Lines 38-39; please define what you mean by 'third-level educated?' Updated manuscript to clarify this is College/University
Pg 9 Lines 42; 21% of participants suggested they were inconvenienced. Did you get any data as to why the felt this way? Do you have data as to how many of these who suggested they were inconvenienced still puckered up and used the tool, or did all of these not use the tool? Have updated manuscript, thanks
Pg 9 Line 55-56, so only 108 of the 140 who completed the feedback survey actually utilised the online check-in? Think this needs to be made clear upfront. Have updated manuscript, thanks
Pg 9/10 Lines 59-5; The 69% who received a call from the COVID-19 team, was this data collected via self-report from the online questionnaire or was it collected objectively through cross-reference with call logs? Online questionnaire
Pg 10 Lines 17-18; those that suggested the quality of the service was 8.5 (n=108), wouldn't they others who did not use the tool also be able to comment on the quality of the service as they were still receiving text messages and had the capability to make/receive phone calls? Have updated manuscript to reflect both populations, thanks

DISCUSSION
Pg 10 Lines 54-55; the statistic suggesting that the online check in tool had good uptake/response rate is a little misleading, as yes 72% used the tool at least once, but they had an opportunity to use it every three days for (unclear how long they received texts for). I think you need to be a little more clear with respect to response rates. The statistic you report is valid, but so is the OVERALL response rate from every text that was sent vs. how many of them prompted a login to the online tool to report on symptoms. Have updated manuscript, thanks
Limitations – while the authors do a sound job discussing the potential limitations of the online tool, they need to do more to address the limitations associated with the research and evaluation component of the paper. There was a skewed sample consisting primarily of healthcare workers (for example). While we want to know about the tool and if it worked well/not well, you have presented a lot of data you suggest speaks to the efficacy of the tool. The readers need a thorough understanding of the limitations associated with these data so as to inform their own decisions about efficacy. Have updated manuscript, thanks

Reviewer: 3

Comments to the Author:
Thank you for the opportunity to review this article titled 'The acceptability of and symptom findings from an online symptom check-in tool for acute COVID-19 outpatient follow-up'. The study presents a novel intervention and offers several important insights. In my comments below I explain how the authors could further strengthen the article and develop more impactful implications and insights from the data:

Existing literature: There needs to be some acknowledgement of the growing body of research on telehealth given the centrality of telecommunications technology in the study, for example, see the following article: Leite, H., Gruber, T. and Hodgkinson, I.R., 2020. Flattening the infection curve–understanding the role of telehealth in managing COVID-19. Leadership in Health Services, 33(2), 221-226. Many thanks for this reference. I have read the publication and have updated the manuscript, thanks

Data audit: How were the contact details for each patient accessed? There needs to be a statement on the protocol followed at a minimum. Since 90% of respondents were healthcare workers it would be useful to reflect on what this bias in the sample may mean for any implications drawn. For instance, to generalise the implications across the general public would be misleading as the respondents will have a much greater level of knowledge and understanding concerning COVID-19. Have updated manuscript to reflect that initials/DOB/symptom onset was cross referenced with hospital electronic records to identify patients, thanks

Follow-up feedback survey: The authors state a feedback survey was used to evaluate acceptability of the online check-in, but this was not sent until 6 months after the SMS 'intervention'. There needs to be a justification for the time lapse as it might be expected that time passed had influenced the perceptions of the participating patients. Have updated manuscript, thanks

Findings: the current presentation of findings is highly descriptive and, therefore, offers limited insights into the role of the online tool. A suggestion would be to segment the patients into groups, possibly based on age bands, healthcare role, etc. and statistically examine whether significant differences exist between groups based on their emotional responses to the tool (i.e. reassuring, helpful, worrying, etc.) and whether significant differences exit between groups based on the perceived usefulness of the tool, satisfaction with the tool, quality of the service, and so forth. Currently, the authors only appear to do this for the healthcare worker group and non-healthcare worker group. Without these more fine-grained insights, the study is less impactful and insightful. Thank you for this comment, however I fear that the group sizes may be too small to adequately compare

In the discussion of the findings the authors state "consideration should be given to including symptoms such as fatigue, headache, myalgia, nasal congestion and pharyngitis to the testing criteria." Again, it would be useful to see how these symptoms varied between patient groups and whether certain groups reported greater relative levels of these symptoms. In its current form, these insights are somewhat lost within the text and are not clearly presented in the results.

Thank you for this comment, however I fear that the group sizes may be too small to adequately compare

Reviewer: 1
Competing interests of Reviewer: None to be declared.

Reviewer: 2
Competing interests of Reviewer: I have no competing interests to declare.

Reviewer: 3
Competing interests of Reviewer: None.

VERSION 2 – REVIEW

REVIEWER	Hodgkinson , Ian R. Loughborough University, School of Business and Economics
REVIEW RETURNED	12-Jul-2021

GENERAL COMMENTS	The authors have done a good job in responding to the comments of the review team. Well done.